Interactive Visual Analytics of Coastal Oceanographic Simulation data

George, Richard; Robins, Peter; Davies, Alan; Roberts, Jonathan C.

Published: 14/10/2012

Peer reviewed version

Interactive Visual Analytics of Coastal Oceanographic Simulation data

Richard L. S. F. George*
School of Computer Science
Bangor University

Peter Robins†
School of Ocean Sciences
Bangor University

Alan G. Davies‡
School of Ocean Sciences
Bangor University

Jonathan C. Roberts§
School of Computer Science
Bangor University

Abstract

Oceanographic simulations typically generate large volumes of unstructured, multi-resolution and spatiotemporal data. Current datasets contain gigabytes of data per time step and include multiple variables. As the complexity and size of the dataset increases, so do the need for tools that can help with the analysis. To this goal we are developing Vinca, (Visualization environment for coastal analysis) a multiple coordinated view system that utilizes a parallel coordinate plot (PCP) view to display all the data variables, which is linked to other temporal and spatial views.

Keywords: Visual Analytics, oceanographic visualization, parallel coordinates

1 Introduction

Coastal shelf studies have assumed greater importance in recent years. In the debate about climate change it is generally agreed that this is resulting in rise of sea level. With 50% of the world's population living within 60 kilometres of the shoreline [3] this sea level rise is a significant issue. Consequently, research into the physical processes influencing oceans and continental shelves must also attempt to predict the likely effect of sea level rise. This research would support the understanding of near shore sea level rise, management of areas increasingly prone to flooding and aid coastal zone engineering. This leads to the need to explore several potential scenarios of different parameter runs. Yet despite this need, coastal shelf researchers themselves acknowledge that forecasting the evolution of the shape and form of coastlines is still not well developed [3] and that historical rates of changes to sea level may not be a reliable guide to their future changes [2].

2 Problem Description

The need to explore different scenarios and high resolution creates a big data challenge. The size of the datasets will only continue to grow, as ocean scientists wish to develop new models with increased complexity, resolution and variables, and correlate them with other datasets. Although it is currently possible to create high-quality visualizations of the data [6], the visualization processes are laborious and non-interactive. Our solution is to develop Vinca (Visualization environment for coastal analysis) using Processing, Java and the ControlP5 libraries. Vinca is a multiple coordinated view system that utilizes a parallel coordinate plot (PCP) view to display all the data variables, which is linked to other temporal and spatial views (see Figure 1).

Ocean scientists wish to drill-down and query specific parts of the data; to obtain quantitative measurements and compare values at different locations and times. Whilst static visualizations may provide insight into oceanographic data, interactive methods provide additional understanding of the domain. Our goal is to investigate how these huge datasets can be interactively explored on a desktop PC.

Our approach is to use several techniques, each designed to overcome a specific problem. A traditional PCP would be unusable, because of the huge overplotting and slow redraw speed, due from the large quantity of data points. Therefore we hierarchically aggregate and simplify the data points to enable fast exploration of the data. This hierarchy is matched to the unstructured grid of the oceanography dataset. We group axis of the PCP, such that the user can elide different axes when necessary. Finally we utilize image caching to increase the performance of the interaction. This work has been developed as a collaboration between oceanographers and computer scientists.

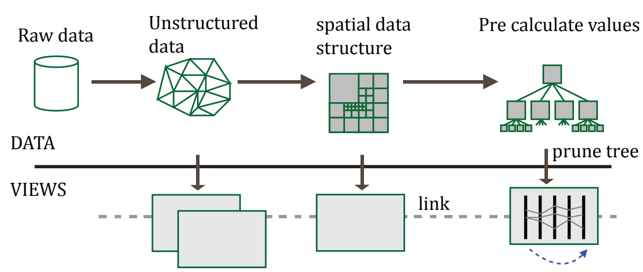

Figure 2: The overall process of Vinca, where we pre-calculate different types of data structures and enable the user to explore through multiple linked views.

3 The data-processing and analytic processes

We calculate depth-average values, and calculate the simulation on an unstructured grid calculating the data using Telemac (opentelemac.org) and visualize it through different views. A schematic of the process is shown in Figure 2. We pre-calculate much of the data, and create several data structures. These are then viewed in multiple coordinated views. The views are linked together through a view-bus configuration, which can permit coordinate between any views. The users can then change how the data is filtered as well as being displayed. Each of the views are framed such that they can be moved around. We create a bespoke windowing management system, such to encapsulate the whole tool in one Processing window. We use the OpenGL renderer to enable fast drawing of the graphical elements. The use of Processing in this way, enables us to migrate the system to different operating environments and will enable us to place it on the web for remote use, for instance.

One of our models is an estuary dataset. We use this to investigate sediment transport and flooding events over the estuary. We use an unstructured grid and store higher resolution points along and in the estuary. In fact, the higher resolution areas extend wider than is traditionally displayed on a map such to calculate over the whole flood-plane. Our runs are predictive, because we wish to explore flooding events and sediment transport over a 100 year period.

We simulate from real-world mean values, and have modelled extreme tidal/fluvial scenarios (worse case scenarios) up to 100

*e-mail: r.l.george@bangor.ac.uk
†e-mail: p.robins@bangor.ac.uk
‡e-mail: a.g.davies@bangor.ac.uk
§e-mail: j.c.roberts@bangor.ac.uk

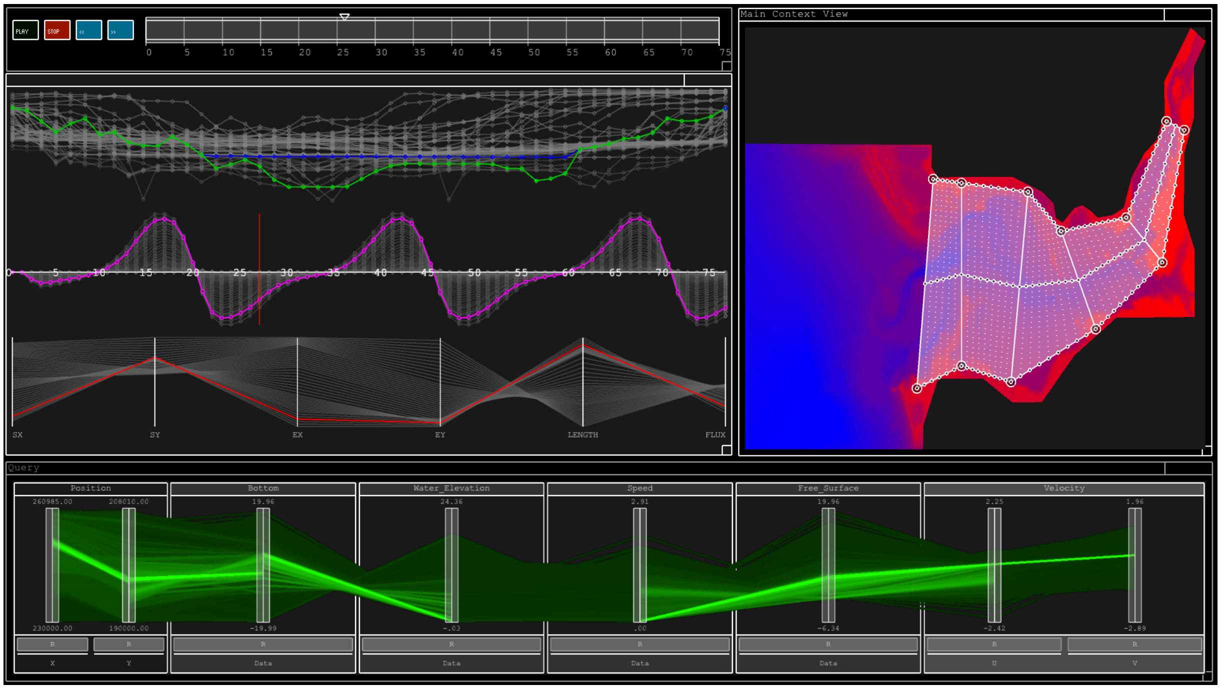

Figure 1: This picture depicts a snapshot from Vinca (VIsualization environment for coastal analysis). This is a multiple coordinated view system that has been developed using Java and the Processing.org libraries. It is designed for the exploration of unstructured oceanographic datasets. We generate the simulation data using Telemac (opentelemac.org) and visualize it through different views. Different views include: time graphs of different variables and profiles (upper left) and the parallel coordinate view (bottom). The parallel coordinate plot is rendered using a frequency binning and outlier preservation algorithm.

years from now. For instance, one of our study estuaries has a railway line on the flood plane. Our hydrodynamic models enable features such as sandbanks, scour pits and tidal channels to be visualized how they change position over time.

To interact and explore the estuary model we utilize several associated data structures. First we load and store the unstructured data, then link these datapoints to a tree (we have implemented both a quadtree and kd-tree, but the oceanographers prefer the quadtree variant, because we also link a plot of the quadtree that is used for selection; and the structure of the quadtree is clearer and easier to understand). We plot all the data in the PCP including the spatial dimensions (x,y and depth). This enables us to aggregate the data in the PCP based on the LOD of the data, rather than the frequency of the data on the PCP.

Typically, two strategies have been used by researchers to reduce the amount of points, and to provide a rendering that is representative of the data. (1) Visual space, and (2) data space [1]. Visual space algorithms such as the screen-space method by Johansson and Cooper [5] provide a visual quality metric, while frequency binning techniques convey the density of the underlying data [4]. In Vinca we have explored different binning techniques, and we have implemented a frequency based binning technique. We base our algorithm on the frequency binning and outlier preservation algorithm of Novotnỳ and Hauser [7].

We draw everything to an off-screen buffer; this is stored as a texture whilst maintaining the alpha channel of the texture. When the user requests an animation (animating through time) we cache all the instances on the first run. With large datasets this does take longer; but it does mean that once this operation is complete the whole data can be animated extremely fast.

4 Conclusions

Work on Vinca is ongoing, it has currently been used to visualize several different estuaries. It enables users to explore the data, to use the PCP to filter the data based on space and value.

The use of the multiple strategies has enabled us to operate the big datasets on a powerful desktop computer. In fact, without these strategies, especially the hierarchical pruning of the quadtree, it would not have been possible to explore different scenarios. The use of Processing also enables the visualizations to be saved to a pdf file, which provides high-quality screenshots to be created.

References

[1] E. Bertini, A. Tatu, and D. Keim. Quality metrics in high-dimensional data visualization: An overview and systematization. *Visualization and Computer Graphics, IEEE Transactions on*, 17(12):2203–2212, dec. 2011.

[2] I. Brown, S. Jude, S. Koukoulas, R. Nicholls, M. Dickson, and M. Walkden. Dynamic simulation and visualisation of coastal erosion. *Computers, environment and urban systems*, 30(6):840–860, 2006.

[3] A. G. Davies and P. D. Thorne. Advances in the study of moving sediments and evolving seabeds. *Surveys in Geophysics*, 29:1–36, 2008.

[4] Z. Geng, Z. Peng, R. S.Laramee, W. Rick, and J. C. Roberts. Angular histograms: Frequency-based visualizations for large, high dimensional data. *Visualization and Computer Graphics, IEEE Transactions on*, 17(12):2572 –2580, dec. 2011.

[5] J. Johansson and M. Cooper. A screen space quality method for data abstraction. *Computer Graphics Forum*, 27(3):1039–1046, 2008.

[6] D. R. Lipsa, R. S. Laramee, S. J. Cox, J. C. Roberts, R. Walker, M. A. Borkin, and H. Pfister. Visualization for the physical sciences. *Computer Graphics Forum*, Online August 2012.

[7] M. Novotnỳ and H. Hauser. Outlier-preserving focus+ context visualization in parallel coordinates. *Visualization and Computer Graphics, IEEE Transactions on*, 12(5):893–900, 2006.